DOCUMENT RESUME

ED 353 925 HE 026 186

AUTHOR Gnagey, William J.; Ostrowski, Kimberly
TITLE The Effects of Changing the Criteria for STAD Team Awards at Midterm: Two Parallel Experiments.
PUB DATE [92]
NOTE 15p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Academic Achievement; *College Students; *Cooperative Learning; *Heterogeneous Grouping; Higher Education; Peer Teaching; *Scores; *Student Evaluation; Teaching Methods; Testing; Test Interpretation
IDENTIFIERS *Student Teams Achievement Divisions

ABSTRACT

This study examined the effects of two variations of the Student Teams Achievement Divisions (STAD) on university students. STAD is a system of cooperative learning which features small heterogeneous teams of 4-6 members who tutor each other on the material in the course and prepare each other for weekly quizzes. Students take the quizzes individually but are awarded bonus points on the basis of the team's mean performance. Rather than awarding bonus points to the teams with the highest average quiz scores which worked in favor of teams with more bright, motivated students, this study looked at the effect of awarding mean improvement points to those teams that showed the highest average improvement scores. Bonus points were awarded for highest average score for the first half of the semester and bonus points for highest average improvement were awarded in the second half. Participants were 158 students in 2 undergraduate psychology courses. Results under the two systems found that using mean improvement points instead of mean quiz scores to award team bonus points was not a robust enough variable to materially affect either the students' achievement or their evaluation of the course. (Also includes seven tables, and a course rating scale.) (JB)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

THE EFFECTS OF CHANGING THE CRITERIA FOR STAD TEAM AWARDS AT MIDTERM: TWO PARALLEL EXPERIMENTS

William J. Gnagey and Kimberly Ostrowski
Illinois State University

ED353925

HE026186

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

William J. Gnagey

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

BEST COPY AVAILABLE

THE EFFECTS OF CHANGING THE CRITERIA FOR STAD TEAM AWARDS AT MIDTERM: TWO PARALLEL EXPERIMENTS

William J. Gnagey and Kimberly Ostrowski
Illinois State University

Introduction

This experiment is the third in a series of investigations of the effectiveness for university students of Student Teams Achievement Divisions (STAD), one system of cooperative learning. These studies have been based on the research of Robert E. Slavin (1983, 1989, 1990, 1991) who focused primarily on STAD's use in public schools. This system features small heterogeneous teams of 4-6 members who tutor each other on the material in the course and prepare each other for weekly quizzes that measure chapter objectives. Students take the quizzes individually, but are awarded bonus points on the basis of the team's mean performance.

In our first experiment (Gnagey, 1988), we found no achievement differences between educational psychology classes that received bonus awards for team improvement and those who received awards based on individual improvement. The individually awarded students, however, rated the course as significantly more effective on the University Course Rating Scale (UCRS).

In our second experiment (Gnagey & Ostrowski, 1991), it was determined that publicizing individual members' contributions to their teams did not influence their individual achievement as measured by quizzes and major examinations.

The present study compares two variations in the awarding of team bonus points: mean quiz scores and mean improvement points.

Even though students in each class are assigned to eight STAD teams at random, when bonus points are awarded to the three teams with the highest average quiz scores, an advantage is given to those who happen to have more than their share of bright, highly motivated students. These teams tend to win the awards over and over again, quiz after quiz. In other words, the bulk of the awards go to those students who are already higher achievers.

We wondered if, beginning at midterm, bonus points were awarded to the three teams with the highest average improvement scores, lower achieving teams would obtain awards more often since they had more room to improve.

We conjectured that reinforcing these lower achieving teams more often during the second half of the semester would increase their level of motivation, their achievement of the course objectives and their evaluation of the course. Since the higher achieving teams would already be doing well, we hoped the new award structure would not affect them adversely.

Hypotheses

The objectives of the present experiments concern the effects of two different criteria for awarding bonus points to collaborative learning teams: mean raw scores and mean improvement scores. The two major experimental hypotheses tested are:

1. Compared to awarding bonus points for mean team quiz scores, awards based on mean team improvement points will result in greater achievement as measured by quizzes and examinations.

2. Compared to awarding bonus points for mean team quiz scores, awards based on mean team improvement points will result in higher student evaluations of the course as measured by the UCRS.

Methods and Techniques

Seventy-seven students in two sections of adolescent development and 81 students in two sections of educational psychology served as subjects in the two experiments. The adolescent development classes were composed of juniors and seniors who were either psychology majors or students preparing to be middle school/junior high teachers. The educational psychology students were mostly sophomores taking their first course in a program designed to produce high school teachers.

For the first half of the semester, all four classes were taught according to the usual STAD format. Each Monday, the assigned chapter was introduced and appropriate material was presented by way of lectures, films or videos. Each Wednesday, heterogeneous teams of 4-6 members convened to work on projects which involved the application of the text material to practical situations.

Each Friday class was divided into three activities. At the beginning of the hour, the instructor held a question and answer session in which students could ask for clarification of difficult concepts in the chapter. When there were no more questions, the STAD teams convened to complete a cross-word puzzle which helped them review the key terms in the chapter. When the puzzles were finished, the teams disassembled. Each student then took a 15-item multiple choice chapter quiz. At the end of the Friday class, the appropriate chapter assignment for the following week was written on the chalkboard.

During class on the following Monday before the new chapter was introduced, all quizzes and puzzles were returned, team average scores and awards were posted (Each member of the three teams with the highest average quiz scores received one bonus point.), and a list of overall grades so far appeared on the wall. The STAD cycle then began for the new chapter. At midterm, a 60-item multiple choice examination was administered over the first seven chapters. The UCRS was also administered at this time (See appendix).

The UCRS is composed of twelve items on which students are asked to evaluate all aspects of the course: clarity of objectives, projects and papers, textbook and other assigned readings, in-class activities, quizzes and exams, feedback, interestingness, instruction, grading procedures, grading fairness, amount learned, expected level of performance (standards), and group activities.

Factor analysis using verimax rotation revealed three principal factors (See Table 1.). The first accounted for 31% of the variance and loaded heavily on items D (in-class activities), G (interestingness), H (instruction); J (amount learned), K (expected level of performance), and L (group work).
The second principal factor accounted for 29% of the variance and loaded heavily on items A (clarity of objectives), B (out of class papers and projects), C (text), E (quizzes and exams), and I (fairness of grades).
The third factor accounted for only 11% of the variance and loaded on item F (feedback).

The test-retest coefficient of reliability for the UCRS total score was .81 using 48 students in similar classes with one administration at midterm and the other four weeks later during the finals.

After the midterm examination, t-tests between the means of all the quizzes, the midterm examination, and the UCRS revealed that there were no significant differences between the two adolescent development classes

(See Table 2.) or between the two educational psychology classes (See Table 3.). A coin was then tossed to select one adolescent development class and one educational psychology class which would serve as experimental groups.

On the Monday following the midterm examination, it was explained to both experimental classes that from then on, bonus points would be awarded to the three teams with the highest average improvement scores. It was further explained that a baseline score (the average of all previous quiz scores) had been calculated for each student and that one's improvement points would be figured by subtracting the baseline score from the quiz score. Students were told that the baseline score would remain the same for the remainder of the course. When the next quiz was returned, each student's baseline score appeared along with the score for that quiz.

The two control classes were taught in the same STAD format that had been used before the midterm. At the end of the semester a 60-item multiple choice examination was administered to all four classes covering only those chapters assigned since midterm. The UCRS was also administered for the second time.

Results and Conclusions

For the adolescent development classes, t-tests between the means of the seven quizzes administered after midterm revealed a significant difference in only one case (See Table 4.). For Quiz 15, the experimental class outscored the control class as predicted. However, no significant differences appeared for either the final examination or the University Course Rating Scale. When the two classes were compared on the three sets of UCRS factor scores, no significant differences appeared (See Table 5.).

For the educational psychology classes, t-tests between the means of the seven quizzes administered after midterm revealed a significant difference in only one case (See Table 6.). For Quiz 10, the experimental class outscored the control class as predicted. However, no significant differences appeared for either the final examination or the University Course Rating Scale. When the two classes were compared on the three sets of UCRS factor scores, no significant differences appeared (See Table 7).

In order to make sure that the experimental treatment actually changed the pattern of team awards during the last half of the semester, Spearman's Rho was calculated between the number of awards teams

received before and after midterm. For both control classes, the correlation coefficients were positive (Rho- .43, Rho- .75) indicating that the award patterns were similar. However, for both experimental classes, the correlation coefficients were negative (Rho- -.27, Rho - -.65) indicating that the award patterns did actually change.

It would appear that in general, using mean improvement points instead of mean quiz scores to award team bonus points is not a robust enough variable to materially affect either the students' achievement or their evaluation of the course. Practically, this means that all of the extra calculations necessary to establish baseline scores and ascertain improvement scores for each of many quizzes are not necessary where achievement and the course evaluation are concerned.

Philosophically, since the whole STAD concept is based upon cooperation rather than competition, one might argue that some criterion-referenced method of awarding bonus points would be more appropriate than the norm-referenced methods compared in the study. However, in another investigation of STAD at the college level, Sherman (1986) found no difference in the achievement of educational psychology students that could be attributed to cooperative vs. competitive goal structures.

References

Gnagey, W.J. (1988). College level STAD: Group vs individual reward. Paper read at Midwestern Educational Research Association meeting, Chicago.

Gnagey, W.J. & Ostrowski, K. (1991). STAD in college. Paper read at Midwestern Educational Research Association meeting, Chicago.

Sherman, L.W. (1986). Cooperative vs. competitive educational psychology classrooms: A comparative study. Teaching and Teacher Education, 2, 283-295.

Slavin, R.E. (1983). When does cooperative learning increase student achievement? Psychological Bulletin, 94, 429-445.

Slavin, R.E. (1989). Research on cooperative learning: Consensus and controversy. Educational Leadership, 47, 52-54.

Slavin, R.E. (1990). Cooperative Learning: Theory, Research and Practice. New Jersey: Prentice Hall.

Slavin, R.E. (1991). Synthesis of Research on Cooperative Learning. Educational Leadership, 48, 71-82.

UNIVERSITY COURSE RATING SCALE

___ A. The objectives of this course were clear to me. I knew what I was expected to learn. (1) hardly ever, (2) occasionally, (3) sometimes, (4) frequently, (5) almost always

___ B. The assigned, out-of-class projects, papers, etc., helped me fulfill the course objectives. They assisted me in mastering the appropriate materials and skills. (1) hardly ever, (2) occasionally, (3) sometimes, (4) frequently, (5) almost always

___ C. The textbook and/or other assigned readings helped me fulfill the course objectives. They assisted me in mastering the appropriate materials and skills. (1) hardly ever, (2) occasionally, (3) sometimes, (4) frequently, (5) almost always

___ D. The in-class activities planned by the instructor helped me fulfill the course objectives. The experiences I had in class assisted me in mastering the course materials and skills. (1) hardly ever, (2) occasionally, (3) sometimes, (4) frequently, (5) almost always

___ E. The measurement devices used in this course were accurate indicators of the extent to which I was fulfilling the objectives. They allowed me to show what I learned. (1) hardly ever, (2) occasionally, (3) sometimes, (4) frequently, (5) almost always

___ F. The feedback in this course was adequate. During the semester, I knew how well I was doing. (1) hardly ever, (2) occasionally, (3) sometimes, (4) frequently, (5) almost always

___ G. My interest in the course was kept high enough to motivate me to do good work. I was able to apply myself. (1) hardly ever, (2) occasionally, (3) sometimes, (4) frequently, (5) almost always

___ H. In my opinion, the teaching of this course was: (1) poor, (2) fair, (3) average, (4) good, (5) excellent

___ I. The grades I received for this course were fair. They were what I deserved for what I learned. (1) hardly ever, (2) occasionally, (3) sometimes, (4) frequently, (5) almost always

___ J. Compared with other courses I have taken at this institution, I feel that I learned: (1) much less than in most courses, (2) less than most, (3) about an average amount, (4) more than in most, (5) much more than in most.

___ K. Compared with other courses I have taken in this institution, I feel that the level of performance expected of me was: (1) much lower than most, (2) lower than most, (3) about the same as most, (4) higher than most, (5) much higher than most.

___ L. The <u>group</u> activities carried out in class helped me fulfill the course objectives. My team helped me master the course materials and skills. (1) hardly ever, (2) occasionally, (3) sometimes, (4) frequently, (5) almost always.

Table 1

Rotated Factor Pattern for UCRS2

Item	Factor 1	Factor 2	Factor 3
A	0.07836	0.69193	0.12490
B	0.36320	0.58192	-0.30593
C	0.19670	0.58731	-0.40395
D	0.77268	0.25255	0.06187
E	0.27006	0.71172	0.12786
F	0.10361	0.23254	0.79150
G	0.60411	0.45104	-0.18749
H	0.69302	0.41617	-0.00718
I	0.03367	0.76926	0.19525
J	0.64069	0.39456	-0.32949
K	0.73521	-0.09374	-0.00112
L	0.69043	0.07623	0.10489

Percent of Variance Explained by each Factor

Factor 1	Factor 2	Factor 3
31.31526	29.25415	11.11909

Table 2

Pre-Treatment T-Test Comparison of Quiz, UCRS1, and Midterm Exam Scores of the Adolescent Development Classes

Section	N	Variable	Mean	SD	t	df	p
1	37	Q1	10.3	3.1			
2	40	Q1	10.6	2.0	0.42	75	0.67
1	37	Q2	11.9	1.9			
2	40	Q2	11.7	3.6	0.38	75	0.71
1	37	Q3	11.4	2.2			
2	40	Q3	12.0	3.1	0.85	75	0.39
1	37	Q4	9.2	2.6			
2	40	Q4	9.5	3.0	0.40	75	0.69
1	37	Q5	10.8	2.6			
2	40	Q5	11.6	3.1	1.22	75	0.23
1	37	Q6	11.6	3.0			
2	40	Q6	11.6	2.8	0.05	75	0.96
1	37	Q7	10.6	2.8			
2	40	Q7	11.7	2.6	1.60	75	0.11
1	37	UCRS1	50.5	6.5			
2	40	UCRS1	51.0	5.5	0.35	75	0.72
1	37	MID	50.2	4.6			
2	40	MID	49.9	6.9	0.22	75	0.83

Table 3

Pre-Treatment T-Test Comparison of Quiz, UCRS1, and Midterm Exam Scores of the Educational Psychology Classes

Section	N	Variable	Mean	SD	t	df	p
5	40	Q1	9.6	2.7			
6	41	Q1	9.8	3.5	0.33	79	0.74
5	40	Q2	10.0	3.3			
6	41	Q2	10.1	2.7	0.29	79	0.77
5	40	Q3	11.5	3.1			
6	41	Q3	11.5	3.8	0.07	79	0.94
5	40	Q4	12.1	1.6			
6	41	Q4	11.6	3.7	0.70	79	0.48
5	40	Q5	9.8	2.8			
6	41	Q5	10.2	3.7	0.53	79	0.59
5	40	Q6	10.2	3.1			
6	41	Q6	11.0	2.7	1.31	79	0.19
5	40	Q7	10.8	2.5			
6	41	Q7	10.4	4.0	0.58	79	0.56
5	40	UCRS1	50.0	5.5			
6	41	UCRS1	51.4	5.2	1.23	79	0.22
5	40	MID	41.9	6.4			
6	41	MID	41.9	7.6	0.05	79	0.96

Table 4

Post-Treatment T-Test Comparison of Quiz, UCRS2, and Final Exam Scores of the Experimental and Control Adolescent Development Classes

Section	N	Variable	Mean	SD	t	df	p.
1(E)	37	Q8	11.4	2.2			
2(C)	40	Q8	10.9	2.5	0.87	75	0.39
1(E)	37	Q9	12.3	2.9			
2(C)	40	Q9	12.9	2.5	0.98	75	0.33
1(E)	37	Q10	10.1	2.9			
2(C)	40	Q10	10.4	3.4	0.40	75	0.69
1(E)	37	Q11	11.5	2.9			
2(C)	40	Q11	11.3	2.5	0.35	75	0.73
1(E)	37	Q12	12.1	2.9			
2(C)	40	Q12	12.2	2.3	0.19	75	0.85
1(E)	37	Q13	10.9	3.7			
2(C)	40	Q13	9.3	4.3	1.78	75	0.08
1(E)	37	Q14	11.9	1.9			
2(C)	40	Q14	10.8	3.0	2.03	75	0.05
1(E)	37	UCRS2	49.4	7.2			
2(C)	40	UCRS2	49.8	5.5	0.27	75	0.79
1(E)	37	FIN	45.2	5.1			
2(C)	40	FIN	44.3	8.6	0.57	75	0.57

Table 5

Comparison of 3 Sets of UCRS2 Factor Scores for the Experimental and Control Adolescent Development Classes

Section	N	Variable	Mean	SD	t	df	p.
5(E)	38	Factor 1	-0.12	0.91			
6(C)	41	Factor 1	0.01	1.16	0.56	77	0.58
5(E)	38	Factor 2	-0.28	1.15			
6(C)	41	Factor 2	0.16	1.06	1.73	77	0.09
5(E)	38	Factor 3	-0.32	1.37			
6(C)	41	Factor 3	0.02	0.77	1.35	77	0.17

Table 6

Post-Treatment T-Test Comparison of Quiz, UCRS2, and Final Exam Scores of the Educational Psychology Experimental and Control Classes

Section	N	Variable	Mean	SD	t	df	p.
5(E)	40	Q8	10.1	2.9			
6(C)	41	Q8	10.1	3.4	0.04	79	0.97
5(E)	40	Q9	10.1	2.3			
6(C)	41	Q9	9.2	3.8	1.32	79	0.19
5(E)	40	Q10	11.4	2.9			
6(C)	41	Q10	9.6	4.0	2.21	79	0.03
5(E)	40	Q11	11.2	3.1			
6(C)	41	Q11	10.8	3.8	0.48	79	0.63
5(E)	40	Q12	9.9	4.2			
6(C)	41	Q12	8.2	6.1	1.40	79	0.17
5(E)	40	Q13	10.0	3.8			
6(C)	41	Q13	9.5	4.0	0.62	79	0.54
5(E)	40	Q14	8.5	3.9			
6(C)	41	Q14	9.5	3.7	1.24	79	0.22
5(E)	40	UCRS2	47.1	6.4			
6(C)	41	UCRS2	48.7	10.2	0.84	79	0.41
5(E)	40	FINAL	39.2	5.4			
6(C)	41	FINAL	38.0	10.0	0.66	79	0.51

Table 7

Comparison of 3 Sets of UCRS2 Factor Scores for the Educational Psychology Experimental and Control Classes

Section	N	Variable	Mean	SD	t	df	p.
1(E)	36	Factor 1	0.01	1.07			
2(C)	38	Factor 1	0.10	0.84	0.37	72	0.71
1(E)	36	Factor 2	0.19	0.80			
2(C)	38	Factor 2	0.10	0.92	0.35	72	0.73
1(E)	36	Factor 3	0.12	0.99			
2(C)	38	Factor 3	0.18	0.70	0.29	72	0.77